

VANESSA – A System for Communication between Deaf and Hearing People

JRW GLAUERT^a, R ELLIOTT^a, SJ COX^a, J TRYGGVASON^a, M SHEARD^{b,1},

^a *University of East Anglia, Norwich, UK {jrwg, re, jt}@uea.ac.uk*

^b *Royal National Institute for Deaf People, London, UK*

Abstract. The VANESSA system provides speech-driven assistance for eGovernment and other Transactions in BSL. The Tool allows speech or text input over a limited subject domain to be translated into sign language, which is then signed for clients using an avatar. The avatar animation is generated synthetically. VANESSA also enables text communication by means of computer keyboards and an interface similar to that used by internet chat room software. The system can be easily modified to translate into other spoken languages, and multiple updates to the content are possible.

Keywords. Sign Language, Synthetic Signing, eGovernment, Translation, Avatar

1. Background and Overview

Deaf people experience communication difficulties every day of their lives because the majority of hearing people do not know sign language. When, for example, a Deaf person needs assistance with filling in complicated forms, the information an assistant needs to convey can be complex, and writing it down quite laborious. Moreover, many people who have been profoundly deaf since birth do not have particularly good reading or writing skills. It must be remembered that sign language is their first language, and (in the UK) English is their second. VANESSA (Voice Activated Network Enabled Speech to Sign Assistant) is an attempt to make communication in assistance environments easier for hearing and Deaf people alike.

It was always the intention that VANESSA should build upon the TESSA system [5, 6]. TESSA helped Post Office Clerks to communicate with their Deaf customers. However it had two significant limitations:

- Speech was the only method by which Post Office clerks could activate the system. Although the system's speech recognition was mostly accurate, if it

¹ The work described here is supported by the European Union as part of the eSIGN Project. We are grateful to colleagues in partner institutions for their support.

did not recognise a particular phrase, the clerk would have to resort to pen and paper, or rudimentary signing, since the large majority of clerks did not know any sign language themselves.

- There was no way, beyond gesturing, (frequently unclear) speech, and/or pen and paper, by which Deaf clients could communicate back.

VANESSA, sought to address these limitations by including text as an alternative method of communication, and giving the client the facility to communicate back to the assistant, using either text or button choice selections. This model was agreed following research by the RNID into what Deaf users would like.

The system facilitates communication between assistants and their Deaf clients in Council Information Centres (CIC's) or similar environments. A signing Avatar conveys the signed information in British Sign Language (BSL) on a screen facing the client. A text version is also displayed. The outline system is shown in Figure 1.

Input from the assistant can be via a microphone or it can be typed in as text using an assistant screen interface. The nearest matching phrases the system has available for signing are then displayed for the assistant. The preferred phrase can be selected and signed for the client. If none of the phrases are appropriate, text alone may be sent.

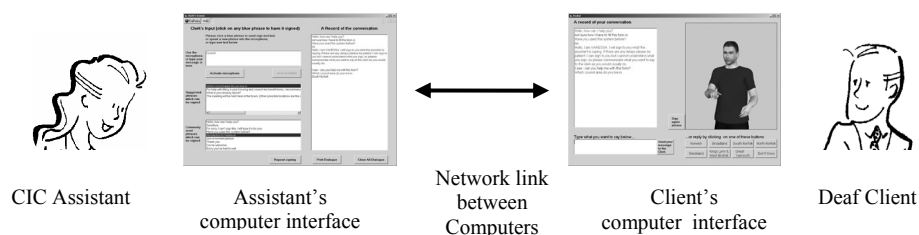


Figure 1 Outline VANESSA system

The assistant's screen also allows the assistant to receive communication back from the client either in the form of typed text, or via a selection of responses, presented to the client as labelled buttons, for them to choose from.

Signs are created synthetically for the system [1,2,3]. The signing has to be accurately described in HamNoSys [4] notation, and this is then transcribed into a form of XML called SiGML[1] which can drive the Avatar.

VANESSA has been developed as a partnership effort between the University of East Anglia (UEA), the Royal National Institute for Deaf People (RNID) and Norfolk County Council (NCC). A working version was installed in the CIC in the Forum in Norwich in August 2004.

A further prototype version of VANESSA has been developed to demonstrate that the system can be easily modified to translate into other spoken languages as well as sign language. There is also a support application that allows new content to be added. If required, such additions may be made in multiple languages.

2. Interface Design

The screen design is the result of discussions between UEA, the RNID and Deaf people. Input and output controls are grouped appropriately and at every possible stage, event messages and guidance are displayed in blue at the top. Boxes and buttons are enabled and disabled according to the current stage of progress, which acts as a guide to what an assistant is able to do at any point in time. The whole screen is disabled, and a message displayed, while the Avatar is signing for a client.

An *Activate microphone* button starts the speech recognition. A short silence terminates recognition. Alternatively, the assistant may type in a phrase. Either input is then matched, using a process which can map the errorful text output by the recogniser to the set of most appropriate signed phrases [5]. The results are displayed and the assistant clicks on any one of them to have it signed. If no suitable phrase is available, the assistant can send the text alone. As keywords are normally enough to identify an appropriate signed phrase, the assistant is encouraged to use this quicker method.

A *Record of the Conversation* box displays a text version of all messages which are sent or received during the course of a conversation. The style is similar to that of an internet chat room interface. Assistant messages are displayed in red and Client messages in blue.

The Client Interface displays an identical *Record of Conversation* box. It also has an input box where a client can type a message, and it displays the *Avatar* which signs all phrases selected by the assistant. When a text-only message has been sent, the Avatar explains this, and indicates the Dialogue Box to its right.

Selection buttons for particular client responses are displayed as labels on the buttons below the Avatar when appropriate. For example, signs for “Which local authority area do you live in?” are accompanied by “Norwich, Broadland, South Norfolk” etc., displayed on buttons at the same time. The Avatar points to the buttons, indicating to the client that they may used to reply. Selected button choices are also displayed as part of the dialogue sequence.

3. Hardware and Software

Two networked computers are required, each with a screen, keyboard and mouse. The assistant also has a microphone. Headset microphones do not encourage use of spoken input as was found with TESSA [5], so the assistant’s microphone is placed next to the their screen on the desk. The assistant’s computer is connected to a printer.

The client’s computer has a high specification to provide good quality graphics for the Avatar rendering. A 2GB Dell Pentium 4, with 512K RAM and nVIDIA GeForce4 graphics card, was used for the installation in Norwich.

VANESSA was written in Visual Basic 6, incorporating ViaVoice modules for the speech recognition, and the SiGMLSigning package [1] which provides the avatar.

A text file is used to maintain hashtable links between a list of phrases available and associated Avatar SiGML files. Alternative ways of saying each phrase are included to build the language model. The SiGML animation files are stored on the Client’s machine to drive the Avatar as required.

A log of all system events is written to a text file, so that use of the system can be monitored. It means the system can be enhanced by adding phrases which are frequently required, but for which no signed version is available.

Output is also possible in the form of a printout of the dialogue from the Assistant's *Record of Conversation* window. This could be useful for dates, times, phone numbers etc., for the client.

4. Content Creation and System Installation

A list of phrases was built up using information from CIC staff and enquiry logs. As it was impractical to include all phrases that could assist with all enquiries, three different groups were selected: a selection of general phrases, phrases to help with filling in forms, and phrases to support dialogue to book a sign language interpreter. The third group of phrases is useful if the assistant or client feels unable to go ahead with the enquiry because the subject matter is too complex.

The accuracy of the translations from English to BSL is key to the success of this type of system. All people involved in the creation and quality verification of the signed animations were expert signers. Translations from English to BSL were first recorded on video. Then, the signs needed to make the phrases were notated using HamNoSys [4], and stored in a BSL lexicon. Editor software [7] was then used to build the signs into phrases and to add in non manual features.

The VANESSA system was installed in the Library at the Forum CIC in Norwich city centre. CIC staff were given comprehensive training on the system

5. Additional Applications

For additional signs to be added, or for a new VANESSA system to be created, a standalone Phrase Maintenance application is provided.

There is also VANESSA Language application, written as a prototype to demonstrate that translation would be possible not only into sign language but into other spoken languages as well. French was selected as the spoken language.

This application requires that a text-to-speech synthesiser is installed. The Euler software was chosen, using the Mbrola speech engine [8], together with a French database for the actual phonemes used.

When French is the selected language, translations are spoken in French, rather than signed. If this application were to be further developed, it is envisaged that talking head technology would replace the signing Avatar on the screen facing the client.

6. Evaluations

To date, there have been three separate evaluations. The salient results are shown in Table 1. All rôle plays were successfully completed, although the quality of avatar signing, the facial expressions and lips in particular, was criticised. This was in part due to the inexperience of the content creation staff; the phrases being the first content they had produced after training. Subsequently the phrases have been reworked, and anecdotal evidence indicates that comprehensibility has improved substantially.

Table 1 Results of VANESSA evaluations

1. CIC evaluation	2 CIC staff and 2 Deaf users. 2 Rôle Plays: Refuse bins not collected, and Benefits form required. All rôle plays successfully completed.
Deaf user comments	VANESSA definitely helped. Better than TESSA. Understood 50% of signing. Text filled in the gaps. Easy to use but more phrases needed. Regional signing differences caused comprehension problems.
CIC Staff comments	Easy to learn and enjoyable to use, particularly finding phrase using keyword. Made enquiry much easier to deal with. More phrases needed.
2. Laboratory evaluation	Hearing member of RNID staff as assistant, 7 Deaf users, 3 Rôle Plays: Broken street light, Benefit enquiry, Interpreter booking. All successfully completed.
Results	Understood signing 60% of time. Used signing in 20/21 tests. VANESSA considered useful by 6/7 participants and enjoyed by 5/7. 4 participants found VANESSA easy or quite easy to use. Layout was liked.
Deaf user comments	I felt in control of the situation. Easy and simple to use. I like the text input, easier than pen and paper, changes possible if you make mistakes. Colour shades on buttons, and larger text input area would be nice.
3. Phrase recognition	15 Deaf participants. 7 representative VANESSA phrases used.
Total phrases recognised	61%. Wide variation in phrase comprehension. Between 2 and 15 participants successfully understood certain phrases
Overall signing quality	Liked: 2, Quite liked: 2, Neutral: 5, Not much liked: 2, Disliked: 4
Participant comments	Harder to understand phrases when out of context. Easier to understand at a slower playing speed. Regional and social differences made comprehension difficult. Repeat playing helped comprehension.

References

- [1] Elliott, R. and Glauert, J.R.W. and Jennings, V. and Kennaway, J.R., "An Overview of the SiGML Notation and SiGMLSigning Software System", *In Fourth International Conference on Language Resources and Evaluation, LREC 2004*, Edited by Streiter, O. and Vettori, C., Lisbon, Portugal, pp. 98-104, 2004
- [2] Kennaway, J.R., "Experience with and requirements for a gesture description language for synthetic animation", *In In 5th International Workshop on Gesture and Sign Language Based Human-Computer Interaction*, vol. LNAI 2915, pp. 300-311, 2003
- [3] Glauert, J.R.W. and Kennaway, J.R. and Elliott, R. and Theobald, B.J., "Virtual Human Signing as Expressive Animation", *In Symposium on Language, Speech and Gesture for Expressive Characters*, University of Leeds, pp. 98-106, 2004
- [4] S. Prillwitz, R. Leven, H. Zienert, T. Hanke, J. Henning, et al. "Hamburg Notation System for Sign Languages - An Introductory Guide.", *International Studies on Sign Language and the Communication of the Deaf*, vol. 5. Institute of German Sign Language and Communication of the Deaf, University of Hamburg, 1989
- [5] Lincoln, M. and Cox, S.J., "A Comparison of Language Processing Techniques for a Constrained Speech Translation System", *In IEEE Conference on Acoustics, Speech and Signal Processing*, Hong Kong, 2003
- [6] Bangham, J.A. and Cox, S.J. and Elliott, R. and Glauert, J.R.W. and Marshall, I. and Rankov, S and Wells, M., "Virtual Signing: Capture, Animation, Storage and Transmission - an overview of the ViSiCAST Project", *In IEE Colloquium on Speech and Language Processing for the Disabled and Elderly*, pp. 6/1-6/7, 2000
- [7] Hanke, T., "Lexical Sign Language Resources: Synergies between Empirical Work and Automatic Language Generation", *Presented at the 4th International Conference on Language Resources and Evaluation, Lisbon*. <http://www.sign-lang.uni-hamburg.de/ilex/HankeLREC2004Poster.pdf>, 2004.
- [8] Dutoit, T., "High-Quality Text-to-Speech Synthesis : an Overview", *Journal of Electrical & Electronics Engineering, Australia: Special Issue on Speech Recognition and Synthesis*, vol. 17 no.1, pp. 25-37, 1997.

Appendix: Screenshots of the VANESSA Client and Assistant Interfaces

